Development and Implementation of a World Wide Web-based Database of Teaching Images

Stephen M. Borowitz M.D., John Jackson, M.S., Venkat Chalasani, Ph.D., Nadine Ellero, M.S.
University of Virginia Health Sciences Center, Charlottesville, Virginia

BACKGROUND. Many of the faculty at teaching institutions maintain slide collections for teaching purposes. These collections are rarely organized or cataloged in any systematic fashion making it difficult for faculty to share these valuable teaching resources with their colleagues. The primary goal of this project was to develop a prototype image database that was accessible across the World Wide Web. This database could ultimately serve as an institution-wide repository for teaching images at the University of Virginia Health Sciences Center.

METHODS. Approximately 600 images from the author's personal collection of clinical slides were sequentially scanned using a Nikon slide scanner and saved as TIFF files of approximately 1.4 MB each. The file size of 1.4 MB was based upon a blind test of pathologists judging when original slides could be distinguished from scanned duplicates. Each scanned image was edited and saved in JPEG format.

While each image was edited and cropped, the authors collected the following data: (1) type of image (i.e. radiograph, photograph, micrograph), (2) free-text description of the image, (3) keyword descriptors, (4) source of image (i.e. owner), and (5) age and gender of the subject (where appropriate).

After an image had been reviewed and cataloged, the record was reviewed by a medical librarian and further indexed by MESH headings. The final record was stored in the MARC record format. The MARC format was employed so that in the future, data can easily be transferred into the Health Sciences Library's information system since the MARC format is utilized by most library information systems.

The Unix® version of Sybase® was chosen as the database repository and all image files reside on a separate Unix-based WWW server. Using computer code supplied by the National Library of Medicine, CGI programs were written to enable WWW-based search and retrieval.

RESULTS. The database currently contains approximately 600 records. Each high resolution image has an accompanying record that includes information on image type, sex, age, MESH headings, owner, department, general description, and a low resolution "thumbnail" image.

The database is accessible throughout the University of Virginia. Access has been restricted to internal users; outside users are blocked from accessing the database with standard http security tools.

Searches of the database are conducted using a simple structured html form. Any or all of the text fields may be searched by text string, title, or MESH heading. A search is specified by manipulating one or more of the available fields. Text expressions may contain combinations of simple Boolean expressions, phrases, or truncated strings.

A search terminates with a display summarizing the search expression, the duration of the search, and the number of matches found. The browsing subset is a collection of thumbnail representations of images with descriptive information from the associated catalog record. The user may scroll through the browsing subset, and: (1) click on a thumbnail image and view a larger version of the image, (2) click on a brief descriptive title and view the full catalog record, or (3) mark an image for later retrieval. Images may be marked and extracted into a subset for later review.

CONCLUSION. The goal of this project was to develop a database that would enable faculty to share teaching slides with colleagues and students. Once the system is fully developed, faculty members will be able to save any teaching slide in the shared database. Faculty, residents, and students will be able to use the database as a teaching resource, a study aid, or as a means of preparing presentation materials.

At present, the main focus of this database is imaging, however currently it can also support digital movies, animations, audio (i.e. heart and lung sounds), or any other digital media owned by faculty. Since all records are stored in MARC record format, the data will eventually be integrated into the Health Sciences Center Library's information system.

1091—8280/97/$5.00